

Towards implementation of the FAIR principles in plasma science

Markus M. Becker^{1,*}, Dirk Uhrlandt¹, Detlef Loffhagen¹, Peter Hill², Marina Prenzel³, and Achim von Keudell³

¹Leibniz Institute for Plasma Science and Technology (INP), 17489 Greifswald, Germany

²York Plasma Institute, Department of Physics, University of York (UoY), Heslington, York YO10 5DD, UK

³Ruhr University Bochum (RUB), 44801 Bochum, Germany

*E-mail: markus.becker@inp-greifswald.de

DOI: 10.5281/zenodo.6483610

Summary

The FAIR data principles were proposed as a guideline for those wishing to enhance the reusability of their data by making them findable (F), accessible (A), interoperable (I) and reusable (R) [Wilkinson *et al.*, *Sci. Data* 3:160018 (2016)]. Since then, various activities aiming at implementation of the FAIR principles in different fields of plasma science have been started:

- Within the project QPTDat, INP works together with partner institutions on research data management (RDM) solutions for low-temperature plasma physics. This includes a close collaboration with the CRC 1316 at RUB, where the focus lies on the establishment of data stewards to support RDM in daily practice.
- The international project Fair4Fusion addresses the RDM needs in the field of fusion plasmas, while PlasmaFAIR at UoY strives to improve the quality and sustainability of plasma research software.

FAIR data principles

Metadata and data should be easily **findable** for both humans and computers.

Data need to be **interoperable** with other data, software or workflows for further processing.

F A I R

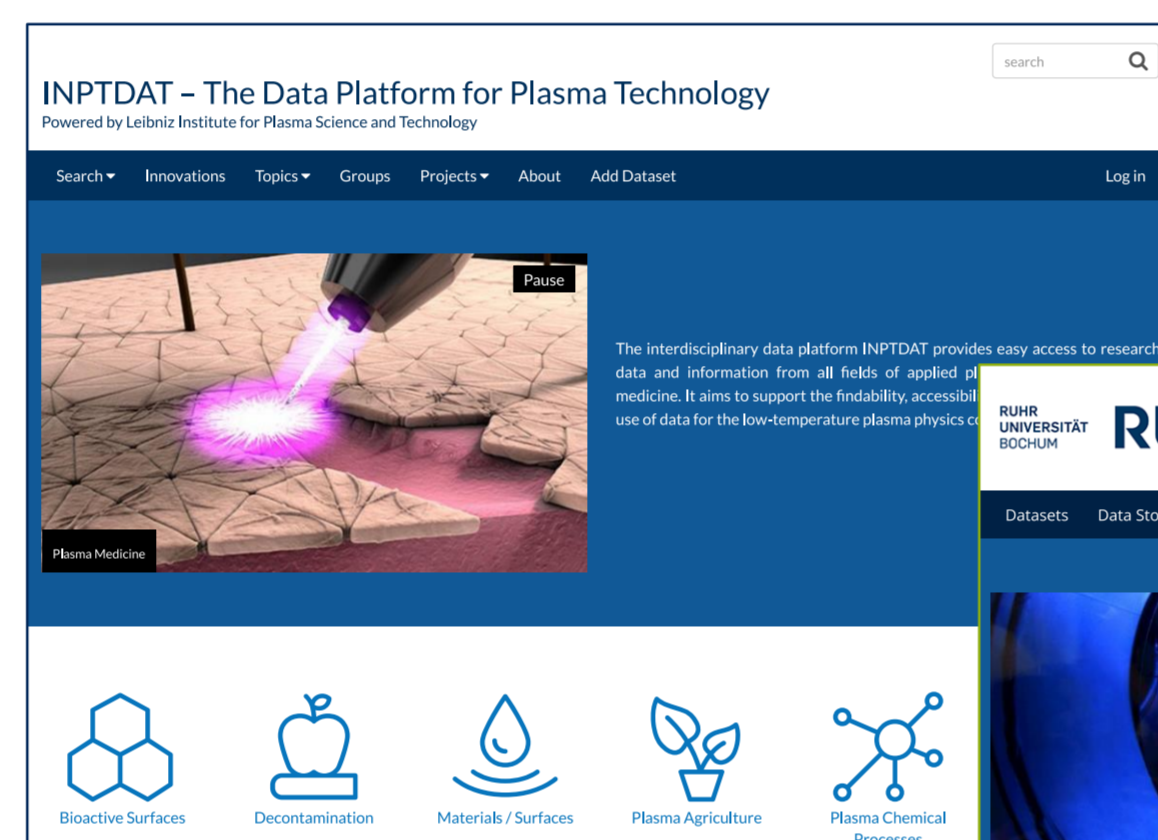
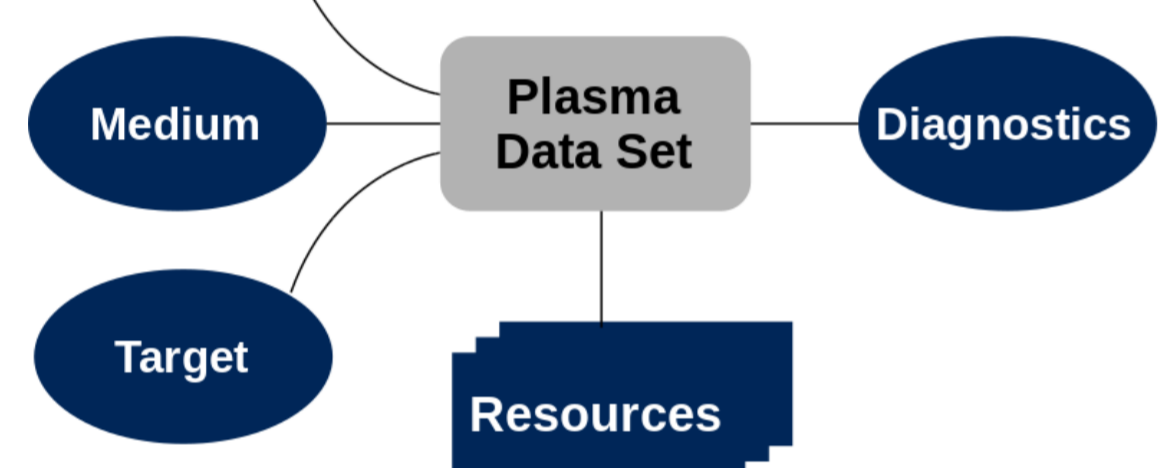
Data must be **accessible**, i.e. users need to know how they can be retrieved.

Metadata and data should be well described so that they are **reusable** within their scope.

Activities in the low-temperature plasma physics community

Plasma Metadata Schema (Plasma-MDS)

St. Franke *et al.*, *Scientific Data* 7 (2020) 439



Research data repositories based on Plasma-MDS



QPTDat
Quality | Plasma Technology | Data

SFB 1316 Transient Atmospheric Plasmas from plasmas to liquids to solids

<https://www.inpdata.de>
<https://rdpcidat.rub.de>
<https://www.plasma-mds.org>
<https://github.com/plasma-mds>
<https://www.forschungsdaten.org/index.php/QPTDat>
<https://sfb1316.rub.de/index.php/en/support-projects/research-data-management>

Aim of QPTDat and the INF project within CRC 1316

- Development of methods and tools for research data management in low-temperature plasma physics
- Repository usage for data sets of peer-reviewed articles, Plasma-MDS for homogeneous metadata
- Highly addressing the community aspects
 - joint working groups for definition of metadata standards and quality criteria
 - activities within CRC 1316 supported by 'data stewards' (representatives from research groups)

INP RUB Meeting

On Friday, 26.11.2021, scientists from Greifswald and Bochum met in a zoom meeting to discuss about metadata of atmospheric pressure plasmas. From both sides the meeting was led by RDM representatives. The results will be used in the development of metadata standards in plasma physics. The next meeting is on 21.01.2022 on low pressure plasmas.

RESEARCH DATA MANAGEMENT

Data Stewards

The challenge of the INF project's work in CRC 1316 is the diversity of the individual review boards. Here, therefore, it makes sense for communication to take place through representative contacts in the groups. These are informed about new developments in the RDM through regular meetings and at the same time can communicate the needs from the groups to

RESEARCH DATA MANAGEMENT POLICY

Collaborative Research Centre 1316 (CRC 1316) – Research Data Management Policy

Research data are a central output of scientific disciplines. They promote the increase of the scientific knowledge base and are the basis for

Event list research data management

■ Research Data Management - International Exchange of Experience
IT Services RUB, SFB 1280, CRC 1316 (RUB, Uni Oulu)
2022-05-11 at 9:00 to 2022-05-12
Location: RUB
Under the motto 'Research data instead of mining coal', UA Ruhr is organizing an international exchange of experience on data management with researchers working as data stewards at the University of Oulu, Finland.

Activities in the fusion plasma community

Application of the FAIR principles to research software (FAIR4RS principles)

- PlasmaFAIR is a new network of research software engineers that aims to help researchers apply the FAIR principles to their software projects, as well as improve the sustainability of software in the plasma science community.
- According to Stodden [<http://doi.org/10.2139/ssrn.1550193>], the biggest obstacle to sharing code and data is the time needed to clean it up for others.
- PlasmaFAIR aims to take some of that burden off researchers through free "software health-checks". Further information are available from <https://plasmafair.github.io>.

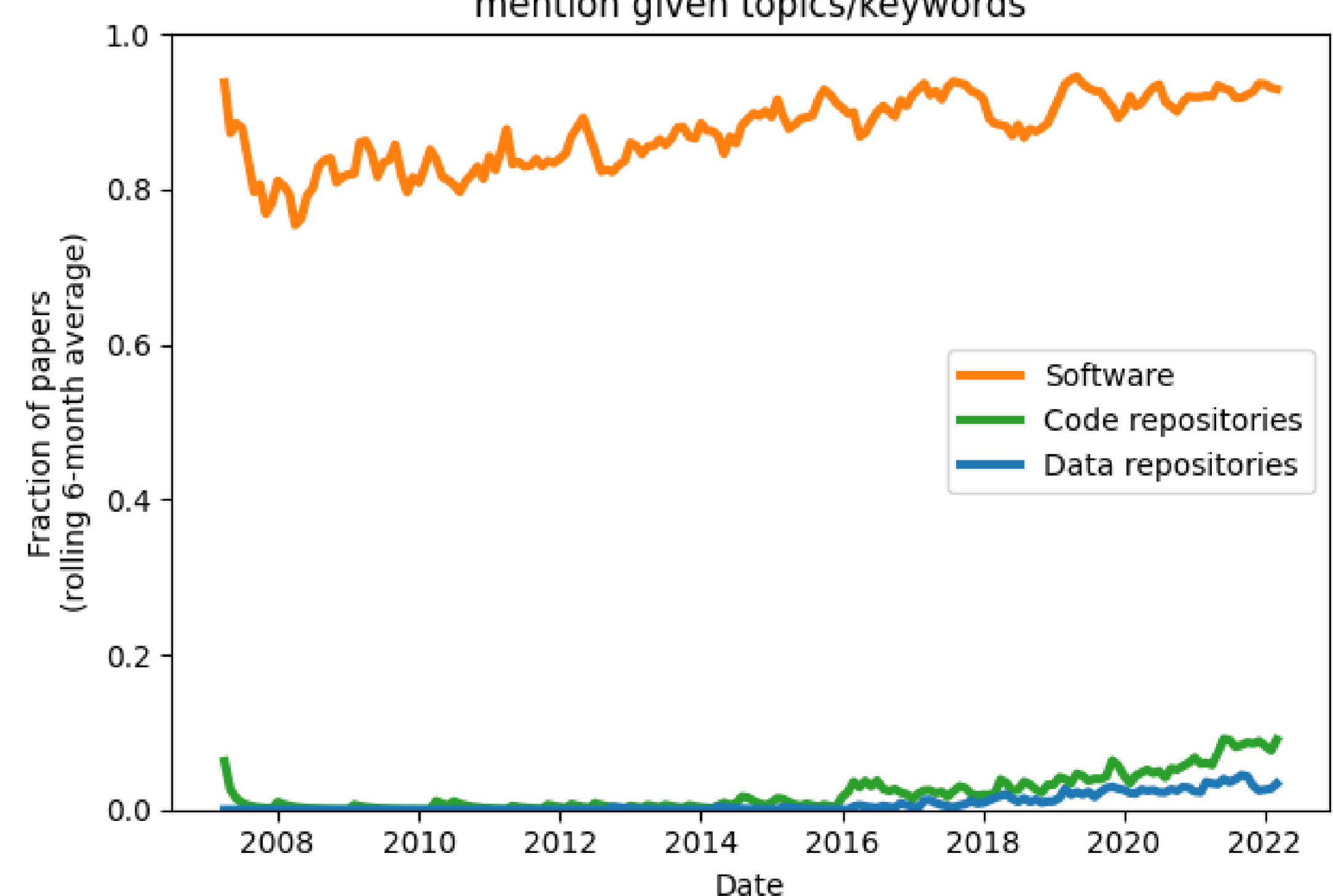
Selected projects that PlasmaFAIR have worked on to date

- FORD: documentation generator used extensively in the Fortran community
- SCENE: tokamak equilibrium solver
- Inference-tools: Bayesian data analysis for tokamaks
- Paramak: parametric 3D tokamak CAD models
- Pyrokinetics: unified gyrokinetic code data analysis



PlasmaFAIR has helped improve build and packaging systems, expand and automate test suites, review and refactor code, and find and fix numerous bugs.

ArXiv plasma physics preprints that mention given topics/keywords



Despite heavy reliance on software, the plasma science community has been slow to embrace FAIR principles, as measured through mentions of code and data repositories in ArXiv preprints.